SPOKES

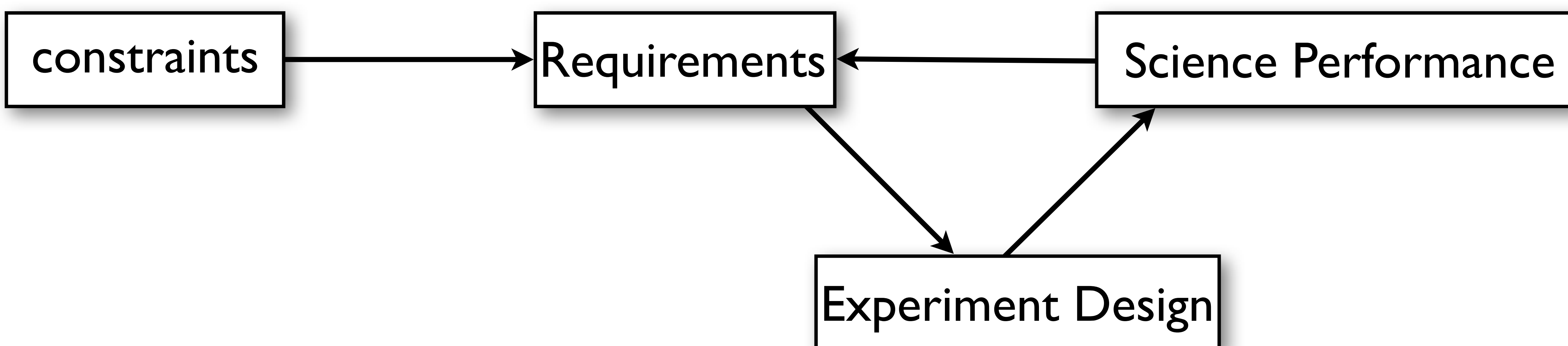
A Facility for End-to-End Simulations of Spectroscopic Surveys

March 5, 2013
MS-DESI Meeting

Brian Nord (FNAL, nord@fnal.gov), *Filipe Abdalla* (UCL), *Adam Amara* (ETH), *Michael Busha* (Zurich), *Oliver Coles* (UCL), *Carlos Cunha* (Stanford), *Tom Diehl* (FNAL), *Brenna Flaugher* (FNAL), *Jaime Forero-Romero* (UAC), *Laurenz Gamper* (ETH), *Lukas Gamper* (ETH), *Ben Hambrecht* (ETH), *Stephanie Jouvel* (Barcelona), *Donnacha Kirk* (UCL), *Rich Kron* (FNAL), *Andrina Nicola* (ETH), *Alexandre Refregier* (ETH), *Will Saunders* (AAO), *Santiago Serrano* (PAU)

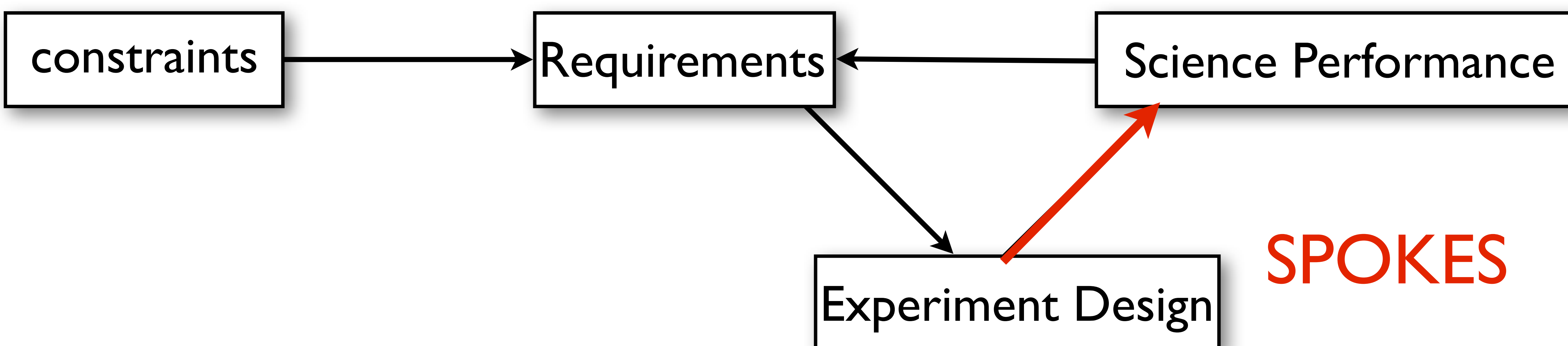
Experiment Design and Development

- Optimize the survey for the science requirements.
- Verify the experiment's feasibility.*
- Prepare to deploy the experiment.



Experiment Design and Development

- Optimize the survey for the science requirements.
- Verify the experiment's feasibility.*
- Prepare to deploy the experiment.



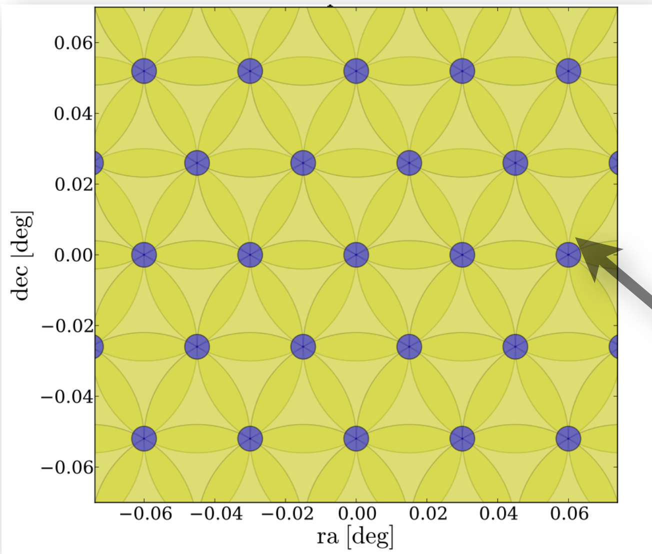
Challenges for a Spectroscopic Survey

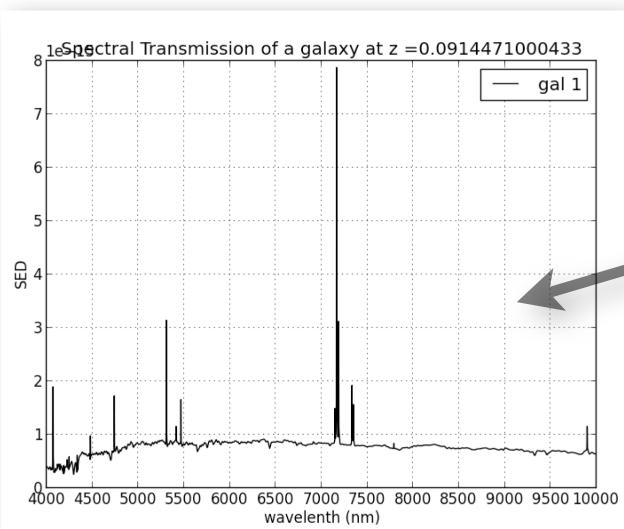
- Requires high precision
- Limited by systematics
- Complex/entangled/non-linear
  [e.g., balance S/N and exposure time]
- Decisions required beforehand
  [e.g., targeting]
- Abundant heritage/foundational work to include.

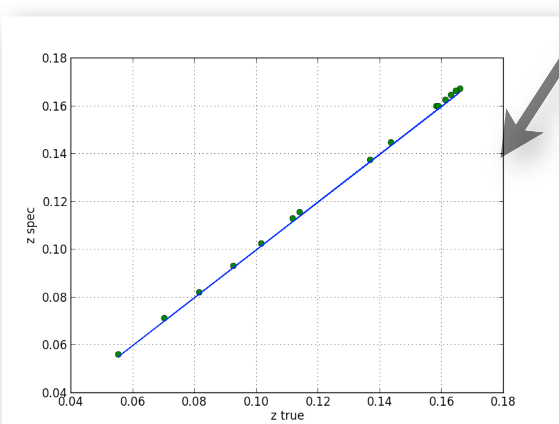
The Key Ingredients of the Framework

• *End-to-End:*	Represent *all* aspects of the experiment with named functions and parameters.
• *Integrated:*	Functions must pass data consistently and clearly.
• *Flexible:*	New functions and new input are simple to ingest and evolve -- also high dynamic range.
• *Fast:*	High-level runs will require less than one day.
• *Reproducible:*	Recover same results for same inputs.

Key Elements of a Spectroscopic Simulation

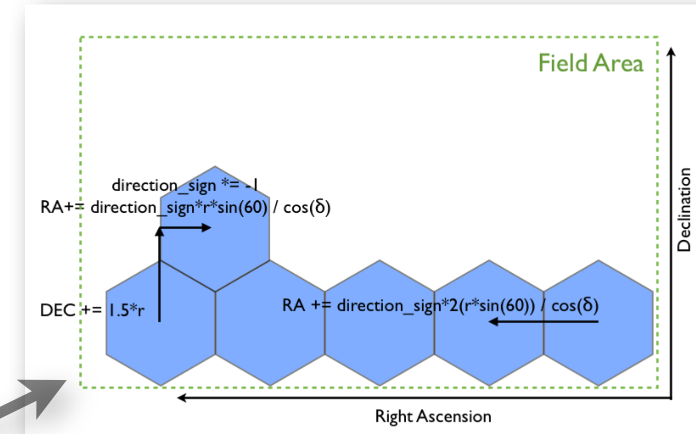


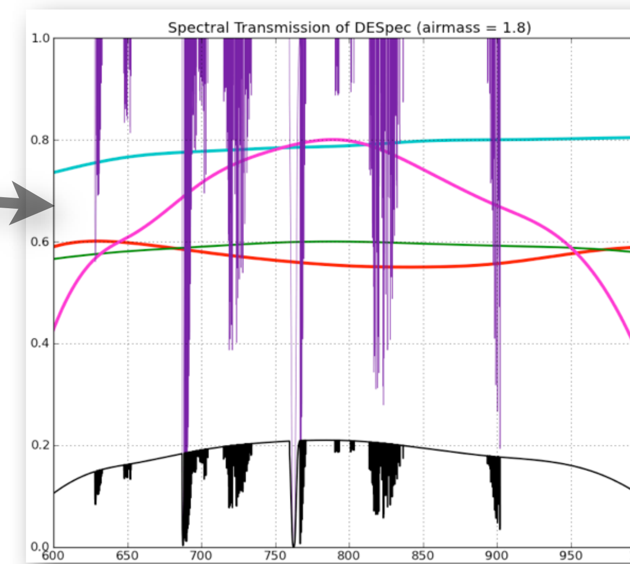


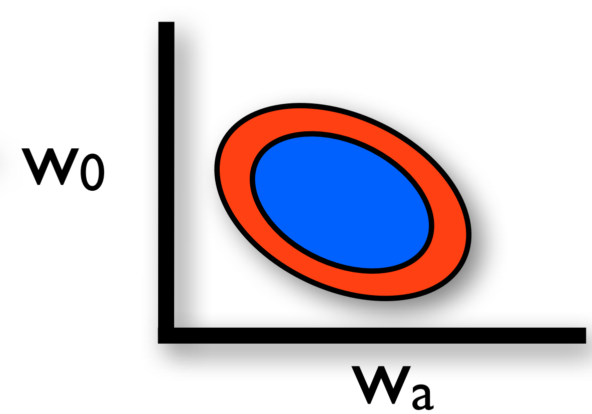


Principle Functions

- Mock galaxies
- Target Selection
- Survey Tiling Strategy
- Fiber Allocation
- Telescope Transmission
- Spectrum Reconstruction
- Noise in spectra
- Spec-Z Measurement
- N(z)
- Selection Function
- Estimate Cosmological Parameters



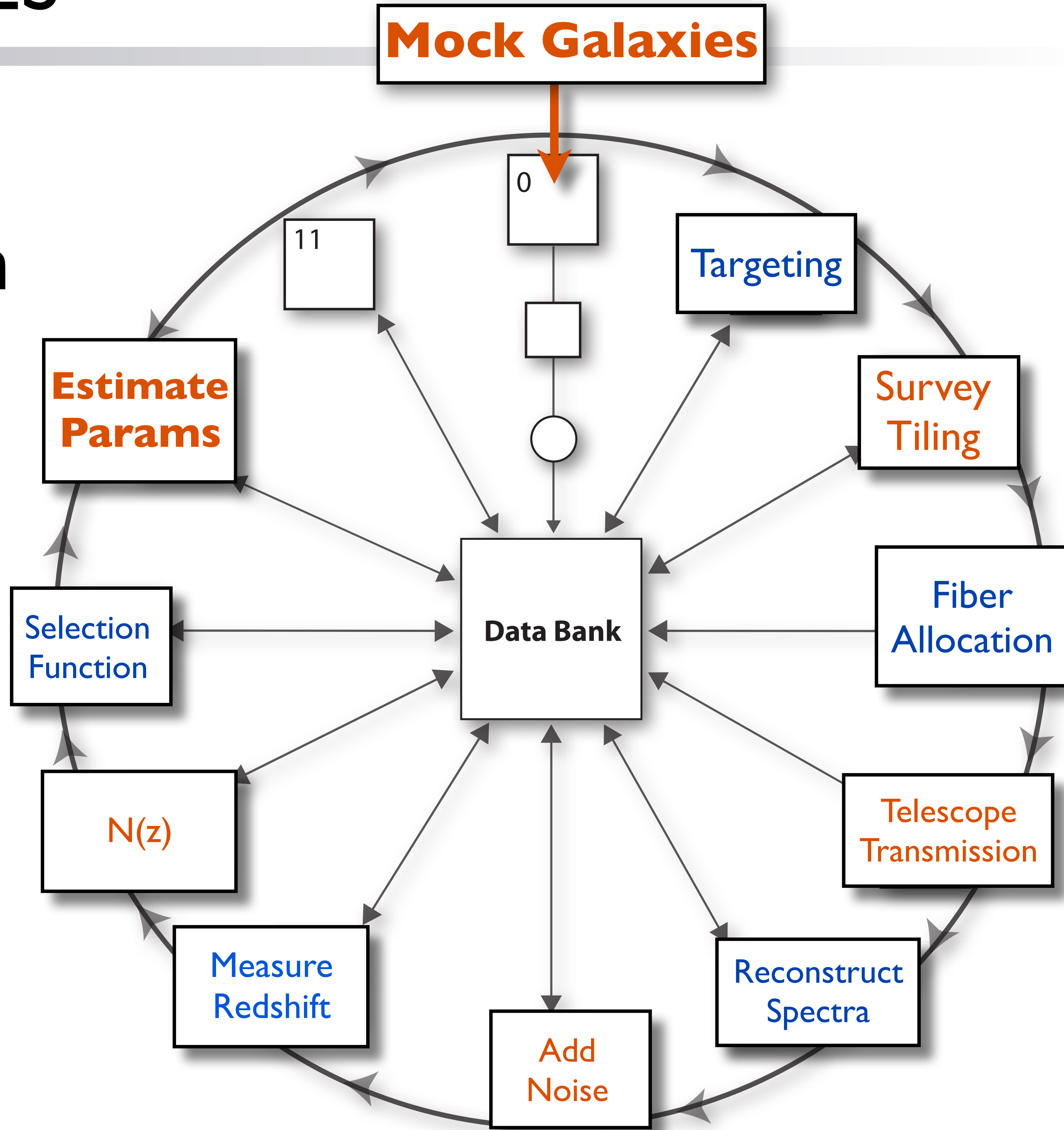




A Solution: SPOKES

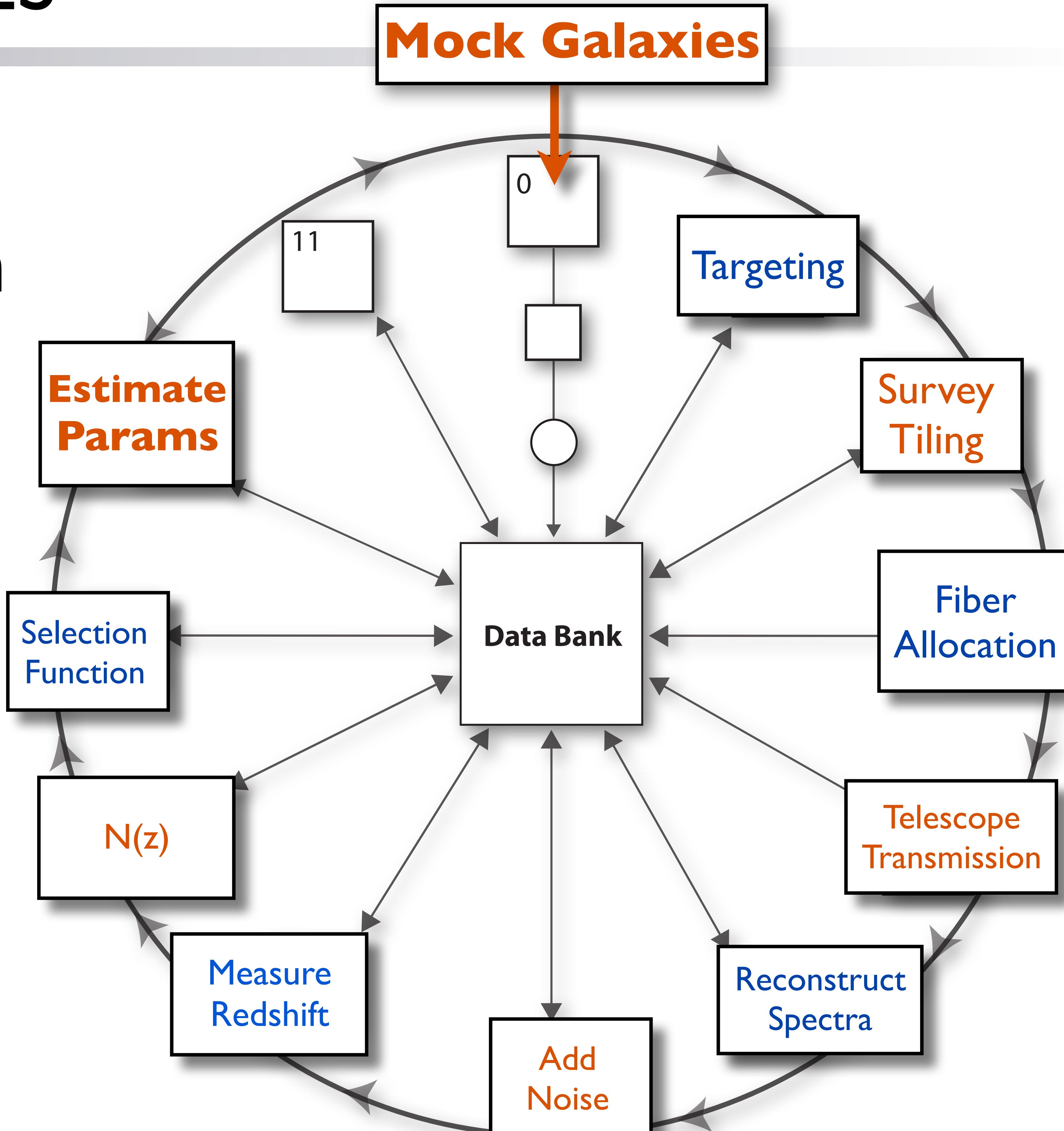
S<u>**P**</u>ectr**O**scopic

KEn **S**imulation



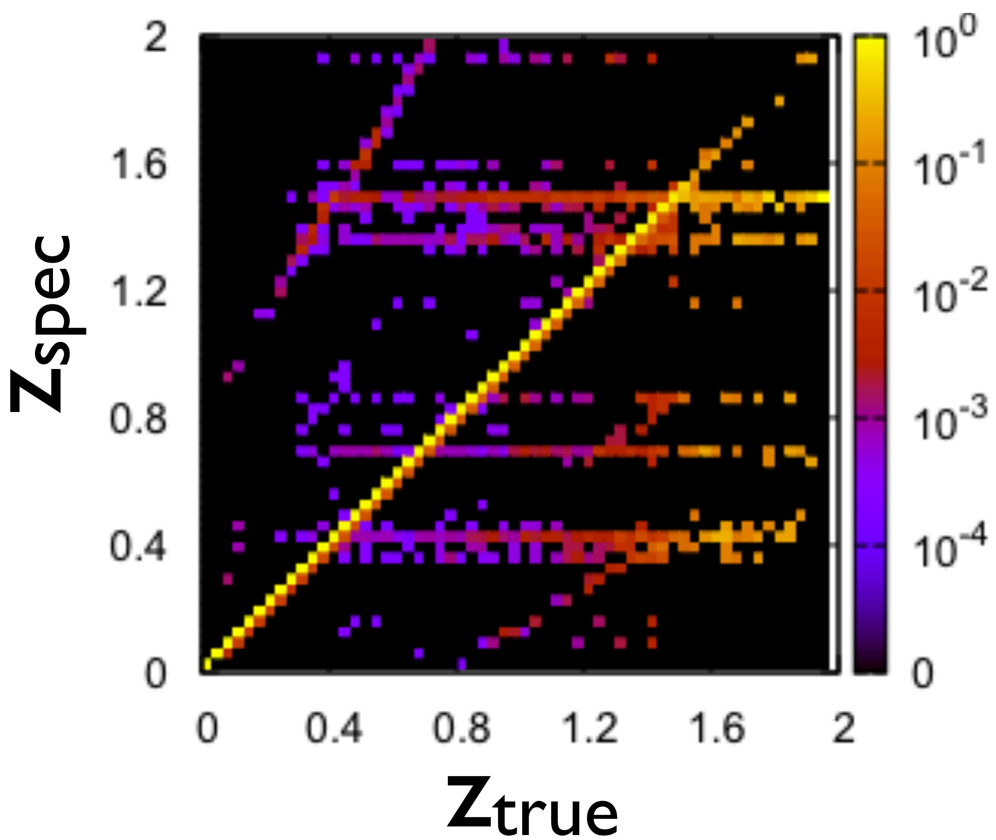
A Solution: SPOKES

SPectr**O**scopic **KE**n **S**imulation

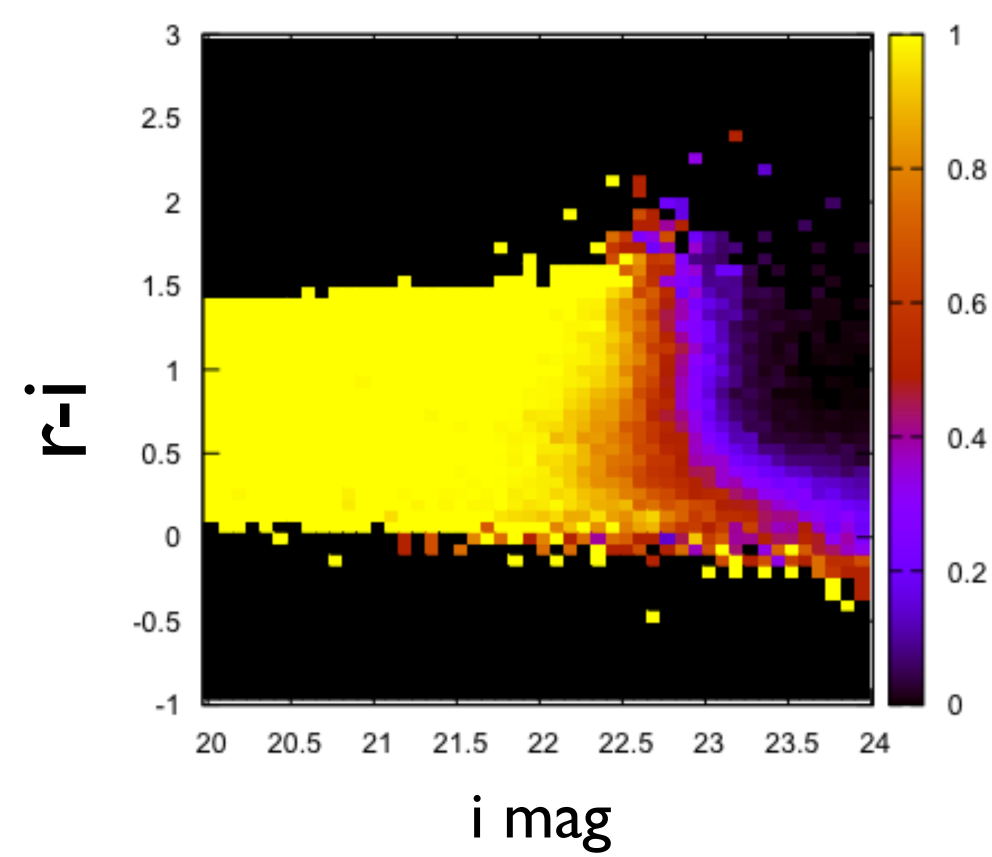


- Architecture integrates modules.
- Infrastructure (API) manages data and documentation with version control.
- Performance Analysis and Benchmarking
- Leverage heritage work.

Example: Noise in Spectra (Cunha et al., 2012)



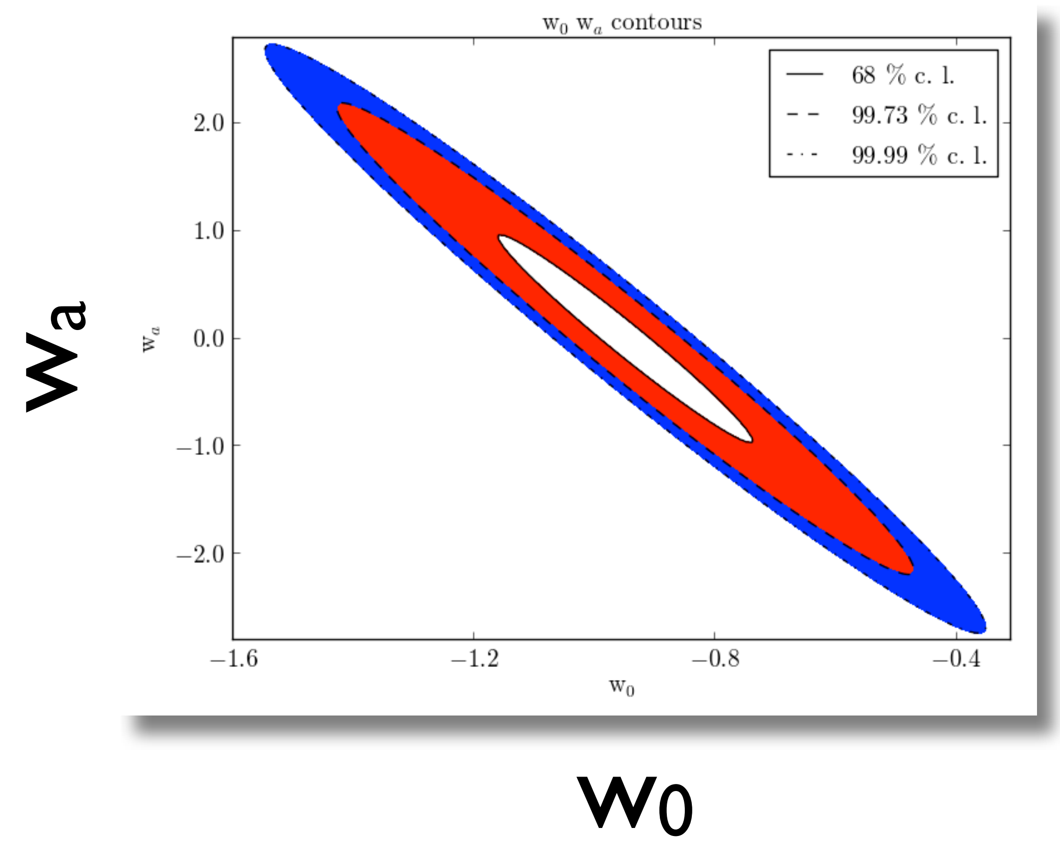
Simulated spectroscopic success

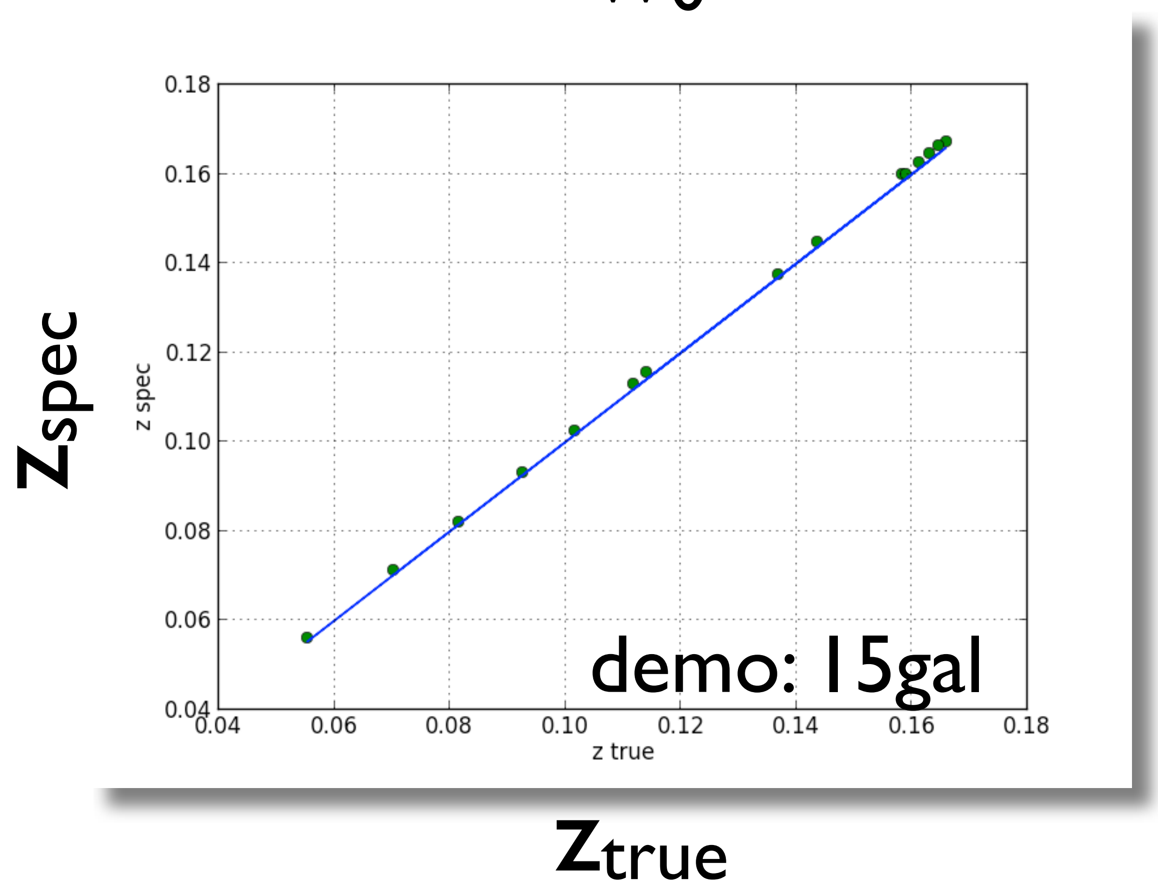


With the simulations, we can test not only completeness, but accuracy.

Preliminary End-to-End Runs

- *initial tests with* DES Mock Catalogs:
  - 50 sq. deg; 7M galaxies
- Assume full region tiled
- 85% fiber completeness
- White noise in spectrum
- Z_{spec} from cross-correlation
- Constraints from FM C_l's





SPOKES in Action: Computational Performance

- Time/Memory usage breakdown
  - Survey Strategy and Spectrum Reconstruction and Redshift measurement take the most time
- Recent runs discovered bottleneck
  - Large memory required to create, store, and manipulate galaxy spectra
  - A solution that *re-balances I/O-CPU Load, preserves code logic and traceability*
    - Function calls will be a new data class in data bank.
    - Parallelize at the galaxy level

Future Work

- Continue integration
- Increase speed
  - optimize current modules
  - implement balancing solution
- Analysis on Identified Priority Issues:
  - Exposure time calculation balance with survey coverage rate
  - Wavelength coverage and science trade-offs
  - dependence on imaging quality for target selection
  - Fiber positioning and fiber density trade-offs

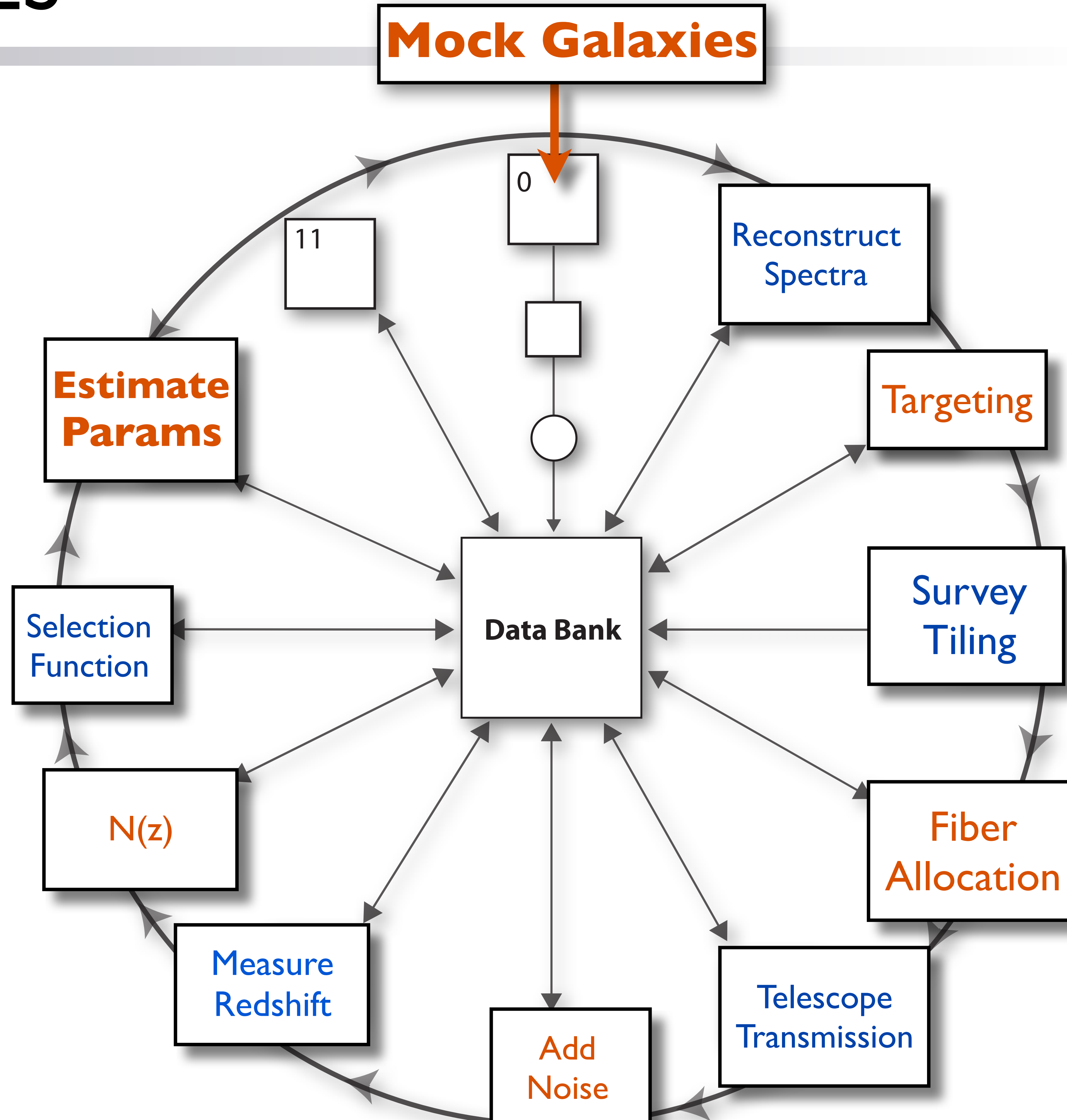
END

Extra Slides

A Solution: SPOKES

End-to-End:

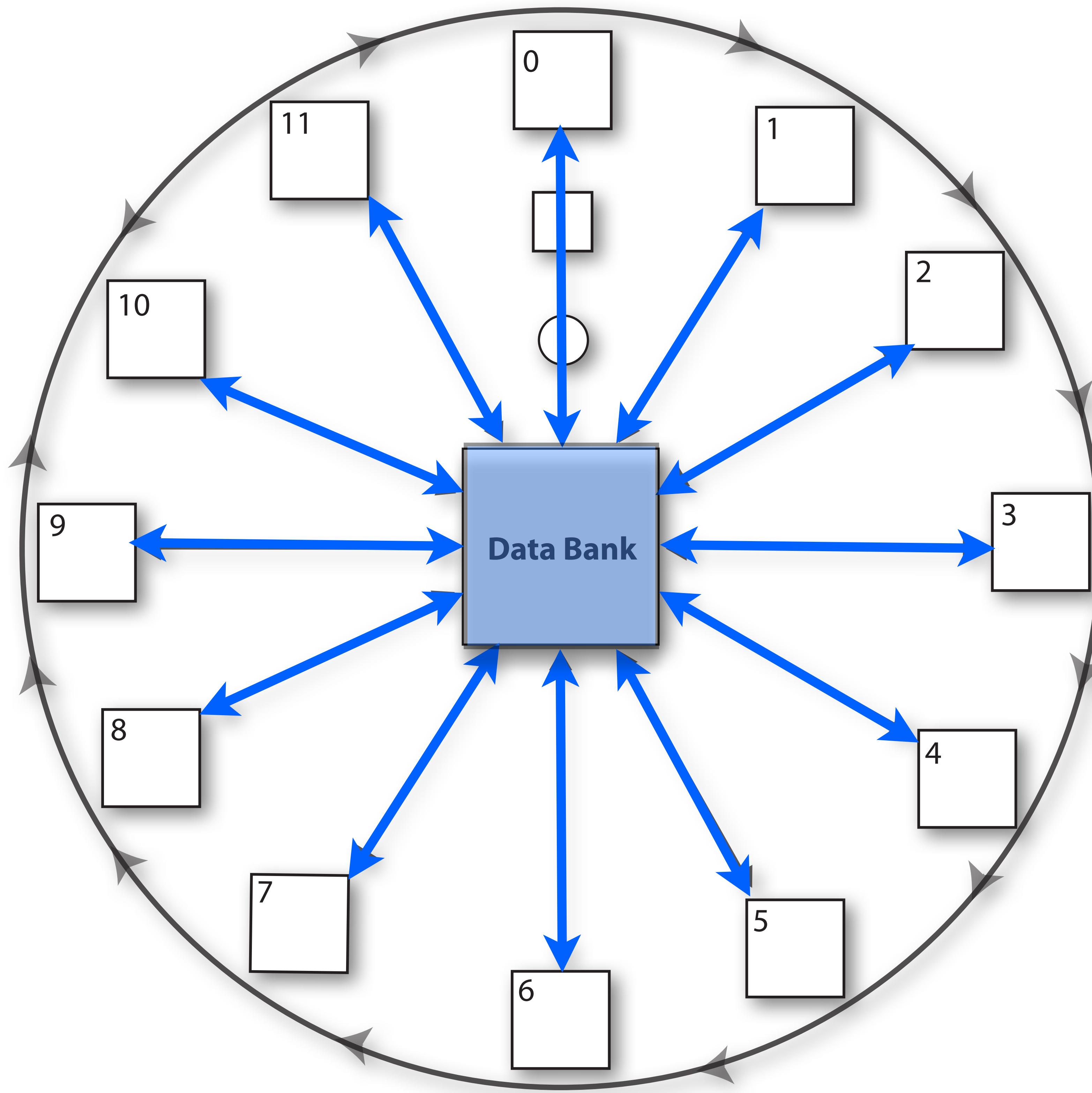
1. From N-body + Mock Galaxies to DE FoM



A Solution: SPOKES

Integrated

1. Each module talks to central data bank for data transfer.
    a. ensures the data are exchanged consistently



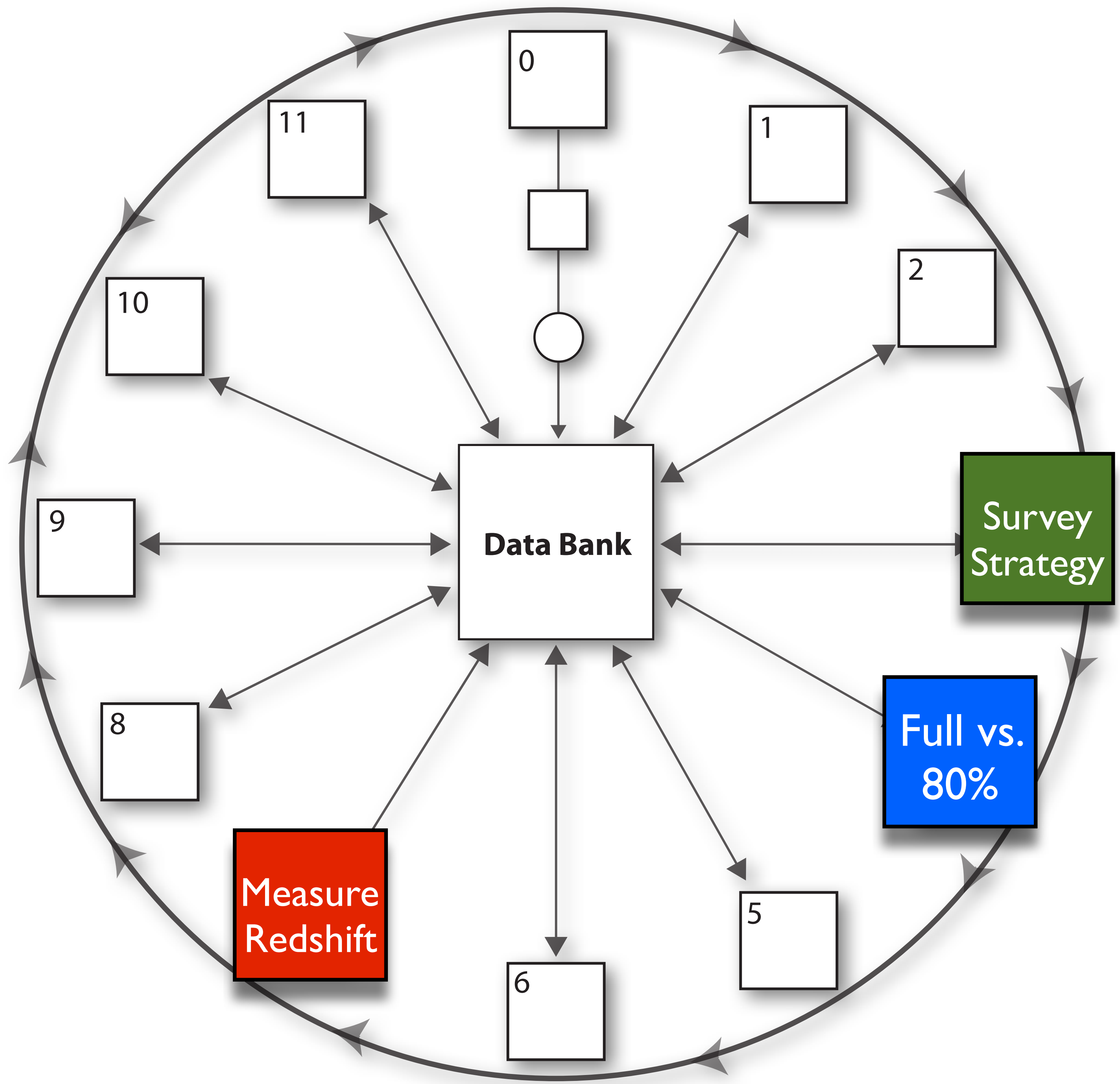
A Solution: SPOKES

Flexible

1. Modules are independent and readily replaceable.
   a) Example: different survey strategy implementations?
   b) already have several ready to go

2. Also, must be dynamic: Representative VS Exhaustive
   a) Example: Run Full Fiber allocation or take prescription of 80%

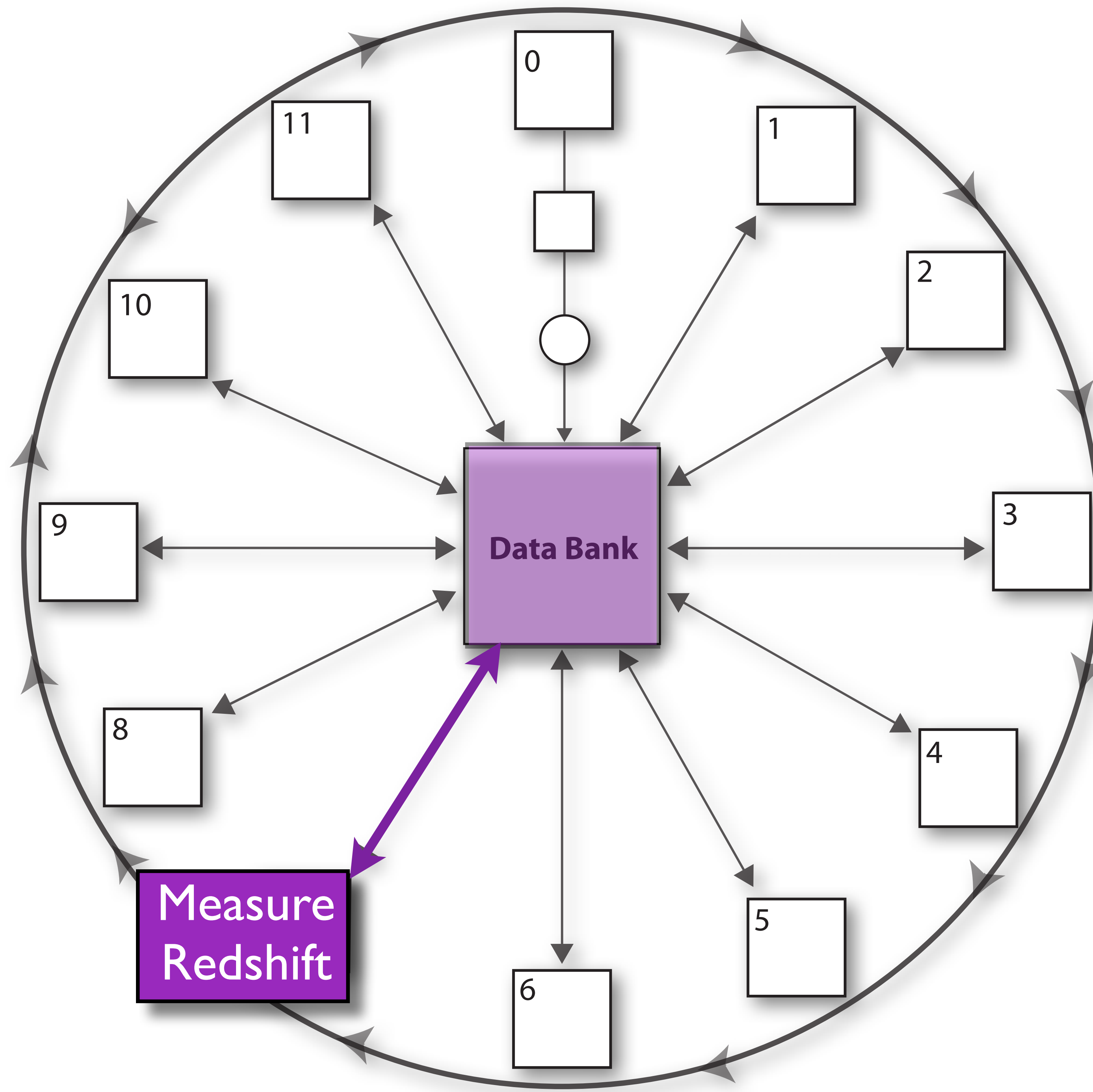
3. Parallelization:
   a) Allows high speed and I/O-CPU load balancing



A Solution: SPOKES

Fast:

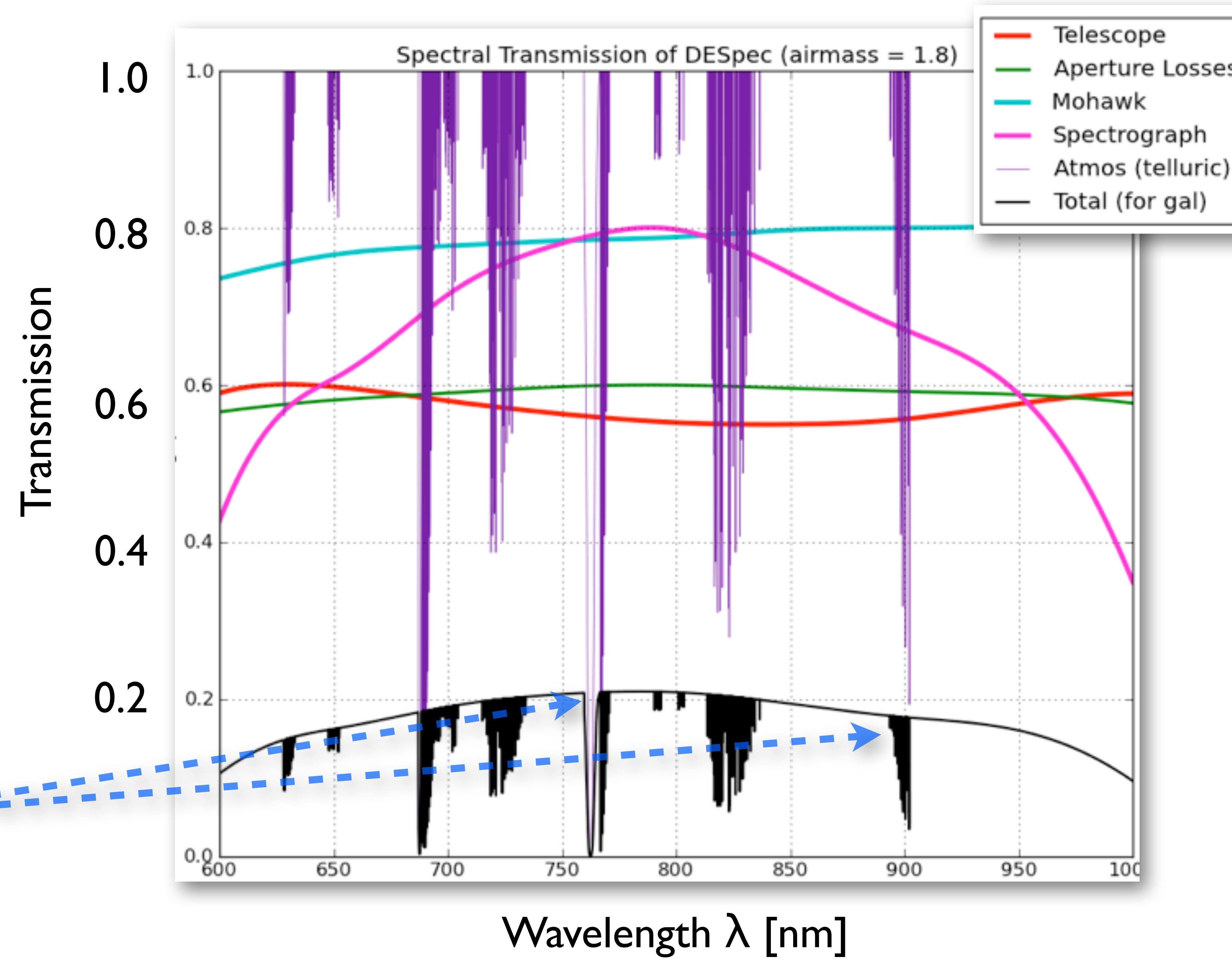
Data access is purely
WYNWYN:
What you need when
you need it.



Example: Telescope Transmission

Detailed level of input:

- Galaxy airmass
- Fiber length
- Apertures
- Connectors
- All optics
- Spectrograph
- Glass types
- Atmospheric absorption with Voigt Profile
- ?



18